

J. Clin. Chem. Clin. Biochem.

Vol. 28, 1990, pp. 399–406

© 1990 Walter de Gruyter & Co.  
Berlin · New York

## The Use of Cluster Analysis in Clinical Chemical Diagnosis of Liver Diseases<sup>1)</sup>

By *Ulrike Folkerts*<sup>1</sup>, *Dorothea Nagel*<sup>2</sup> and *W. Vogt*<sup>1</sup>

<sup>1</sup> *Institut für Klinische Chemie und Laboratoriumsmedizin, Deutsches Herzzentrum München*

<sup>2</sup> *Institut für Klinische Chemie, Klinikum Großhadern der Universität München*

(Received December 5, 1989/March 9, 1990)

**Summary:** Diagnostic judgement is usually based on recognition of patterns. Unfortunately more than three quantitative data cannot be judged simultaneously without help of mathematical methods. Working on laboratory reports, a clinician usually goes linearly through the columns and reduces quantitative to qualitative data. Therefore the medical decision process should be improved if data reduction is performed with the aid of mathematical methods for pattern recognition.

A total of 191 consecutive outpatients with a tentative or proven diagnosis of hepatobiliary disease were examined clinically, clinically chemically and partly histologically. Nineteen clinical chemical parameters were determined. Prior to pattern cognition, a principal component analysis was performed. Using six factors, accounting for 72.4% of total variance, cluster analysis was done, applying a hierarchical algorithm for ascertaining a starting partition, followed by the k-means algorithm. The validity of the solution was scrutinized, and a stable structure was found with nine clusters. Patients with a rejected suspect of liver disease were mainly located in clusters 1, 6 and 7. Cluster 1 also contains patients with compensated cirrhosis without inflammation, idiopathic hyperbilirubinaemia, focal nodular hyperplasia and haemangioma of the liver. In contrast, one third of cirrhotoses, all with inflammatory activity were assigned to cluster 5. Patients with primary biliary disease were distributed among clusters 2, 3 and 4. All malignant neoplasias were assigned to cluster 9. More than 50% of fatty livers were classified to cluster 7. Cluster 2 and 8 contain only one patient with primary biliary cirrhosis (cluster 2) and fatty liver hepatitis (cluster 8). The follow-up of 66 patients also showed clinically meaningful changes of cluster assignment.

### Introduction

Diagnostic judgment in hepatobiliary diseases is often necessarily based on a large battery of laboratory tests. But with an increasing number of laboratory results it becomes more and more impossible to take into account interrelationships between the parameters observed just by intuition. Therefore the use of formal abstract operations could be helpful in pro-

viding the clinician with clinical chemical patterns instead of a lot of single data.

Many attempts have been made to find multivariate concepts (for example 1–4) which should support the clinician in the interpretation of clinical chemical data. Commonly used methods, however, of interpreting clinical chemical data (such as discriminant analysis) have the drawback that they require a correct final diagnosis made independently of the clinical chemical data.

<sup>1)</sup> This publication is part of thesis of U. F.

In our opinion a suitable model for a diagnostic support should meet two requirements:

- to avoid the above-mentioned drawback, the procedure should be independent of the diagnoses and
- the information from a patient's data should be condensed to one total unit of information to facilitate the interpretation of the data.

A more detailed discussion of this problem can be found in l. c. (8).

Several previous publications of our group (5–8) dealt with cluster analysis for defining groups of patients with similar patterns of clinical chemical data. The model based on these results and constructed with patients suffering from thyroid diseases was found to be a meaningful method of data reduction in clinical chemistry, clinically helpful and suitable for increasing diagnostic information.

Readers who are interested in a more comprehensive discussion of pattern cognition and recognition in clinical chemistry are referred to l. c. (8).

In the present study patients suspected of having liver disease were classified by cluster analysis using clinical chemical parameters only. The biochemical entities were determined absolutely independently of premises such as symptoms, signs or diagnoses. An important question was whether it would be possible to find a stable classification in spite of the large number of laboratory tests used, and furthermore whether the grouping would be clinically meaningful.

## Materials and Methods

### Acquisition of data

All 191 outpatients, referred to an expert hepatologist<sup>2)</sup> during an 18-month period with a tentative diagnosis of hepatobiliary disease or follow-up of known liver disease, were included in the present study. They all were examined by the same expert. If necessary, the liver was histologically examined by needle biopsy.

Nineteen clinical chemical parameters were determined quantitatively in each patient (cf. fig. 2). Based on these parameters and all the clinical information, the hepatologist made a final diagnosis for each patient. The case history and the findings of the physical examination were listed in a standardized protocol. Laboratory data were stored and processed without knowledge of clinical data.

### Principal component analysis and cluster analysis

The 19 clinical chemical parameters were standardized to zero mean and unit variance, then a principal component analysis

was performed. The resulting six main components were used as variables for classification. Each patient was represented as a point in a six dimensional space.

As a partitioning clustering method we chose the variance criterion, which minimizes the variances within each group. The algorithm we used is known as the k-means principle (9). For obtaining a preoptimized starting partition we used the hierarchical clustering method of *Ward* (10), which optimizes the same criterion. The choice of these methods implies the use of *Euclidean* distances. The "right" number of clusters was determined by using the concept of cluster validity (8).

### Hardware and programs

All programs were run on a Siemens P 7.570. For principal component analysis program BMDP4M was used. For cluster analysis we used the algorithms by *Späth* (9), while all other programs were written by us.

## Results and Discussion

Certain problems are involved in applying numerical classification procedures, because cluster analysis includes a broad spectrum of mostly empirical methods that lack commonly accepted recommendations for deciding the following questions.

### Choice of clinical chemical parameters

The first problem to be solved was the selection of the kind and number of clinical chemical parameters. Obviously the choice of variables used to describe a patient has great influence on the result of classification. We had to carefully consider medical and mathematical aspects. Thus, the aim of the present study was to support the clinician in differential diagnosis of diseases of the liver, and since this is a most complex organ with many biochemical functions, the application of a large number of clinical chemical parameters seemed to be necessary. Since there are many partly contradictory recommendations for using laboratory tests in hepatology, we selected a combination of 19 quantitative parameters, which, nevertheless, are probably not considered the best by everyone.

### Principal component analysis

From a mathematical point of view the medically motivated use of 19 variables for clustering a sample of 191 patients is not practicable, as the distribution of the patients in a 19-dimensional space would be too rarefied, i. e. a stable classification is improbable. Further, the unwanted, strong influence of linearly dependent parameters on the classification result needs to be diminished.

<sup>2)</sup> Prof. Dr. *Josef Eisenburg*, formerly a member of the medical staff of the Medizin. Klinik II, Klinikum Grosshadern, has taken the clinical part and provided us with diagnoses. We are greatly indebted for his kind support.

Therefore we decided on a principal component analysis with the patients' data. As the method was required only for data reduction before using cluster

analysis, we dispensed with trying a pathophysiological interpretation of the principal components. The criterion for limiting the number of principal components was to extract only components with eigenvalues >1. The six components determined by this method accounted for 72.36% of the total variance of the clinical chemical parameters. The sequence of the factors 1–6 is arranged in decreasing order of variance represented by each factor (fig. 1). Figure 2 shows which proportion of variance of the 19 clinical chemical parameters is explained by the six compo-

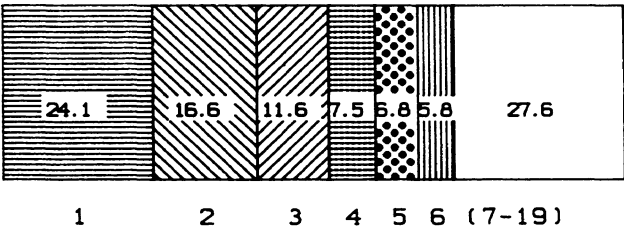


Fig. 1. Variances explained by factors 1–6.

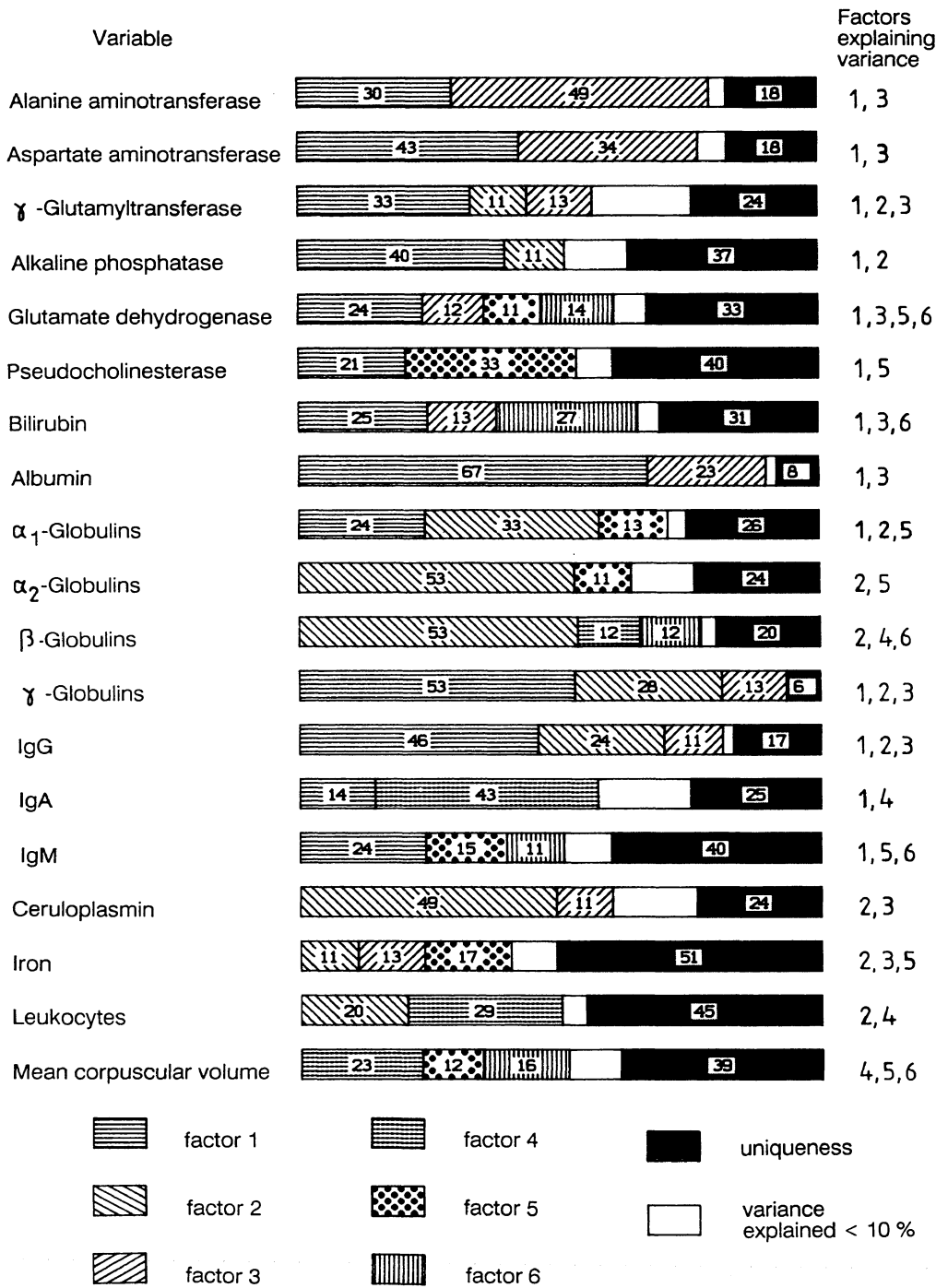


Fig. 2. Variances of the 19 clinical chemical parameters explained by factors 1–6.

nents extracted and which part is not taken into account. As there is a quota of variance that is inexplicable by six factors only (uniqueness), the principal component scores are values which are statistically estimated. That means a loss of information, a drawback which can be accepted for the sake of projecting data from the 19-dimensional to the six-dimensional space.

### Number of clusters

Another problem is to decide on the optimal number of clusters. In making this decision we considered two aspects.

Firstly, from a pragmatic point of view the number of groups depends on the medical purpose of the study and therefore on a certain subjective decision made by the user. The aim of the present work is to find a model supporting the clinician in multivariate interpretation of laboratory results concerning differential diagnosis of liver diseases. Therefore we decided that a differentiating model with about ten or fifteen groups would be more useful than a model with two or three groups only. Secondly, however, the decision should not be made on grounds of subjective considerations only. A procedure recommended e. g. by I. c. (11) for partitioning clustering methods is to plot the criterion value (in our case the sum of variances within each group) against the number of clusters. For 2 to 25 partitions we obtained a monotonously decreasing curve with local minima of the optimization function at the group numbers 4, 9, 12 and 16. The result for group number 9 was most significant.

The problem of choosing the correct number of clusters can also be solved in connection with cluster validity. We assumed that the solution with the highest validity represents the true structure of the data set in the best way. Using this concept, we have a method for separating true cluster structures from artefacts produced by an algorithm (7, 8).

Necessary conditions for the validity of a classification are as follows:

1. The partition obtained will be found for different samples of the same general population (internal stability).
2. Small changes in the input data will produce only small changes in classification (external stability).

### Internal stability

The internal stability of different partitions was measured as follows. Half of the patients of the sample were taken randomly and clustered in the same way

as the original data. The procedure was repeated with 35 different random samples each for the group numbers 2 to 25. As found in prior investigations, a number of 35 repeats is sufficient (8). The results were compared with the original clustering for the selected patients. For calculating the consistency of the different partitions we used "Rand's measure corrected for chance" (12). Figure 3 shows the median values of the 35 stability coefficients plotted against the group numbers 2 to 25. In the interpretation of figure 3, not only the absolute values have to be considered, but above all the form of the stability curve. Considering former results (8), it is supposed that a stability curve created by a data set without cluster structure increases nearly monotonically, so that the existence of peaks or shoulders of a function points to solutions with high stability. There are peaks for the group numbers 9, 11, 18 and 19. The maximum of the curve shows the partition of 9 clusters to be the classification with the best internal stability.

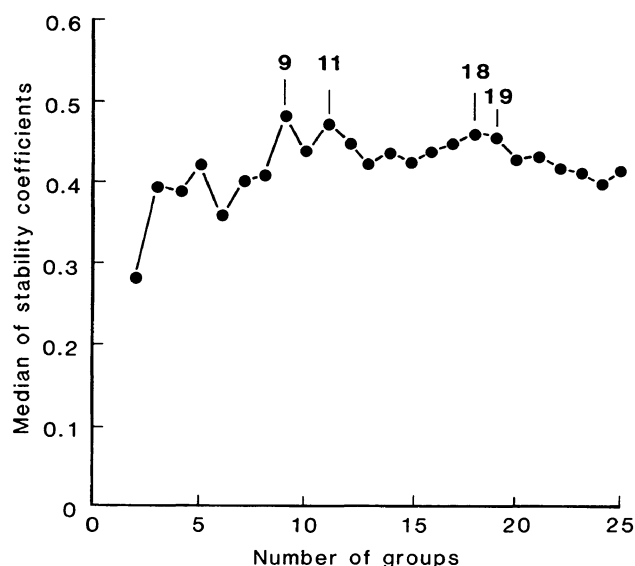


Fig. 3. Internal stability of structure.

Taking into consideration the subjective aspects described above and the results of the two mathematical procedures, the optimal number of clusters is nine.

### External stability

A method of proving the external stability is to assign all the patients with input data (changed by adding errors) to the clusters by a reclassification algorithm. The percentage of correctly reclassified patients is a concrete measure of the external stability of the partition found; alternatively, if it is calculated for each cluster, it is a measure of the validity of a single group. According to the analytical imprecision of the methods, we added 35-times random errors to the

original data. The distribution of the errors was known from the daily routine work. The average reclassification rate was calculated for the total partition, as well as for each single cluster.

With 82% of correct reclassification we regard the external stability to be sufficient. Table 1 shows that the results for the large clusters are not as good as for those with few patients. But it is precisely for the small groups that a good stability is very important from a medical point of view.

Tab. 1. Matrix of relative frequency (%) of assignment by 35 reclassifications.

Group the patient belongs to	Group the patient is assigned to after adding analytical errors										n
	1	2	3	4	5	6	7	8	9		
1	75					18	6		1	74	
2		100								1	
3	4		93	1			2			12	
4				97	3					5	
5	3			1	91	3	2			14	
6	12				1	81	6			32	
7	7					6	87			45	
8								100		1	
9	3					6			91	7	

To give a graphic presentation of the groups for each cluster the mean value of the six variables (principal components) F1 – F6 are shown as vectors, whose length is a measure of the respective value. The total means of the factors lie on the arch of a circle. So the similarity of difference of the groups is clearly visible by their graphic representation (fig. 4). Groups which

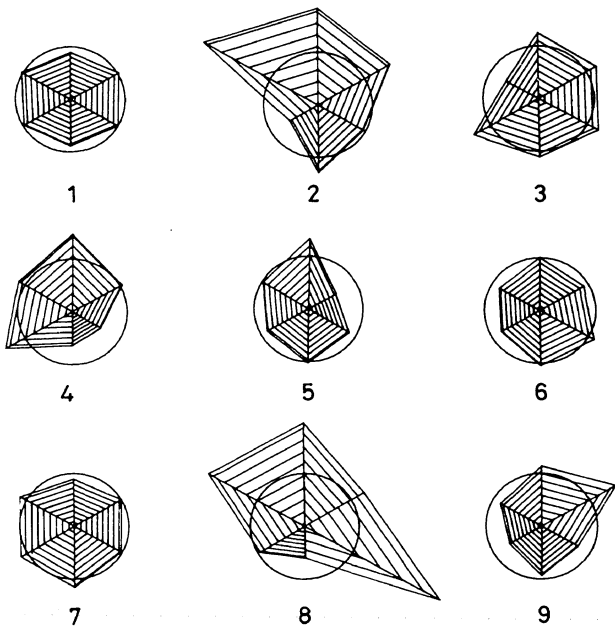


Fig. 4. Graphical patterns of the nine groups. Beginning from noon in clockwise direction: factor 1, 2, 3, 4, 5, 6.

exchange many patients in consequence of fluctuating input data show a similar graphical pattern, e. g. clusters 1, 6 and 7.

The percentage of objects that have changed groups due to imprecision of input data can be used for showing the vicinity of and interrelations between different clusters (fig. 5). Clusters which interchange patients are connected with lines, the numbers above correspond to the vicinity of two clusters given by the average erroneous reclassification rate of a group  $G_i$  to a group  $G_j$  and vice versa.

Clinical interpretation of the clusters found

For checking the clinical meaning of the entities derived only from laboratory data, we compared the clinical chemical classification with the final diagnoses (which are partly based on morphological criteria) and with clinical findings.

It is self-evident that we had not expected to find an exact consistency between a classification based solely on laboratory tests and a classification based mainly on morphological criteria. We also did not intend to replace the previous nosological system by a clinical chemical classification like the present model or to anticipate any kind of diagnosis. But the model is only useful for the clinician if there is some connection between the grouping found and the current diagnostic concepts. We succeeded in finding a relation between the two kinds of classification.

Tables 2 and 3 show the final diagnoses for the sample and their distribution among the clusters. Patients, in which the suspect of liver disease could be rejected, were mainly allocated to cluster 1 (71.4%), to cluster 6 (11.4%) and cluster 7 (14.3%). Cluster 1 also contains the patients with diagnosis of idiopathic hyperbilirubinaemia, focal nodular hyperplasia and haemangioma of liver. Twenty five percent of the patients diagnosed as having cirrhosis are classified in cluster 1. These cirrheses are all well compensated and without inflammatory activity. In contrast to this, the 30% of cirrheses representing the main part of cluster 5 show inflammatory activity, reduced liver function and a history of repeated decompensation. The patients originally given the diagnosis of primary biliary disease are distributed among clusters 2, 3 and 4. All cases of malignant neoplasia (liver cell carcinoma and liver metastases) are assigned to cluster 9. More than half of the patients with steatosis are classified in cluster 7. The two singular groups (2 and 8) each contain a single patient with an extreme form of disease.

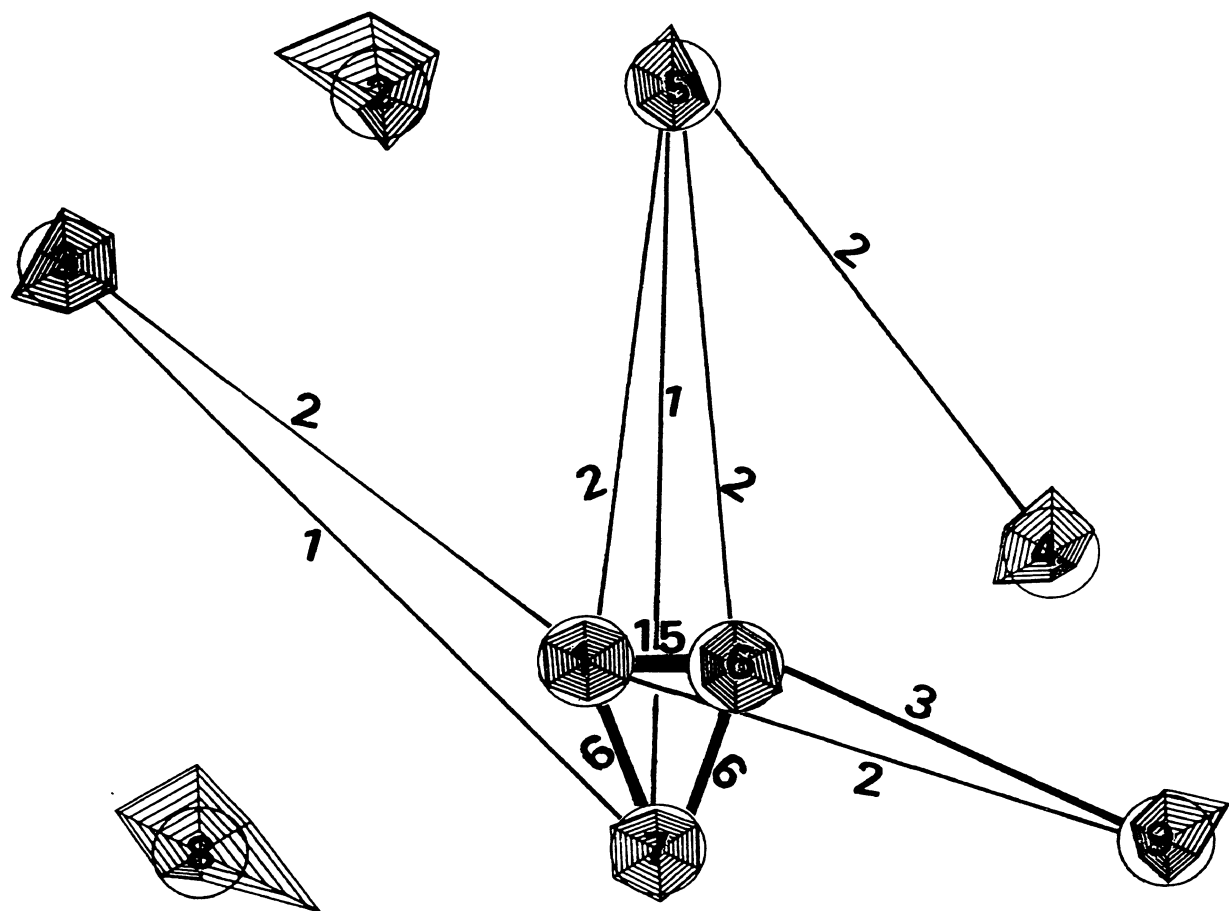


Fig. 5. Illustration of similarity and vicinity of clusters found.

Tab. 2. Final diagnoses (n = 191).

Code	Final diagnoses	Number of patients	
		n	%
1	No liver disease	35	18.3
2	<i>Gilbert's</i> syndrome	3	1.6
3	Chronic hepatic congestion, state after cardiac recompensation	3	1.6
4	Focal nodular hyperplasia	5	2.6
5	Haemangioma	2	1.0
6	Haemochromatosis	4	2.1
7	Cystic tumour	1	0.5
8	Moderate fatty changes	25	13.1
9	Fatty liver	24	12.6
10	Cirrhosis	40	20.9
11	Chronic persistent hepatitis	27	14.2
12	Chronic active hepatitis without cirrhotic transformation	6	3.2
13	Acute viral hepatitis with prolonged course	3	1.6
14	Primary biliary cirrhosis	5	2.6
15	Acute cholecystitis with obstructive liver injury	1	0.5
16	Primary sclerosing cholangitis	1	0.5
17	Acute alcoholic hepatitis	1	0.5
18	Liver cell carcinoma	3	1.6
19	Tumor metastases	1	0.5
20	State after partial resection of the liver due to malignant hepatoma	1	0.5

It is evident from table 3 that all non-singular clusters contain patients of different diagnostic categories, which means that an exact correspondence of the two classifications does not exist. Certainly the main reason for the discrepancy is neither a failure of clustering methods nor an accumulation of diagnostic errors, but the impossibility of completely separating the disease groups by means of clinical chemistry only.

Therefore the model does not provide the clinician with a new concept of diseases or with an improved form of the previous concept. But information on the cluster membership of patients provides the clinician with clinical chemical patterns instead of a mass of values. An experienced clinician is accustomed to work with patterns by evaluating symptoms and signs, and anamnestic or histologic findings. There should therefore be little difficulty in working with laboratory patterns defined by mathematical procedures.

There exists a graphically presentable individual figure for every patient, which can be compared with the figure of the group the patient has been assigned to and with the figures of other groups. In our experience, the monitoring of laboratory results in the

Tab. 3. Clusters versus final diagnoses (for explanation of diagnosis codes see tab. 2)

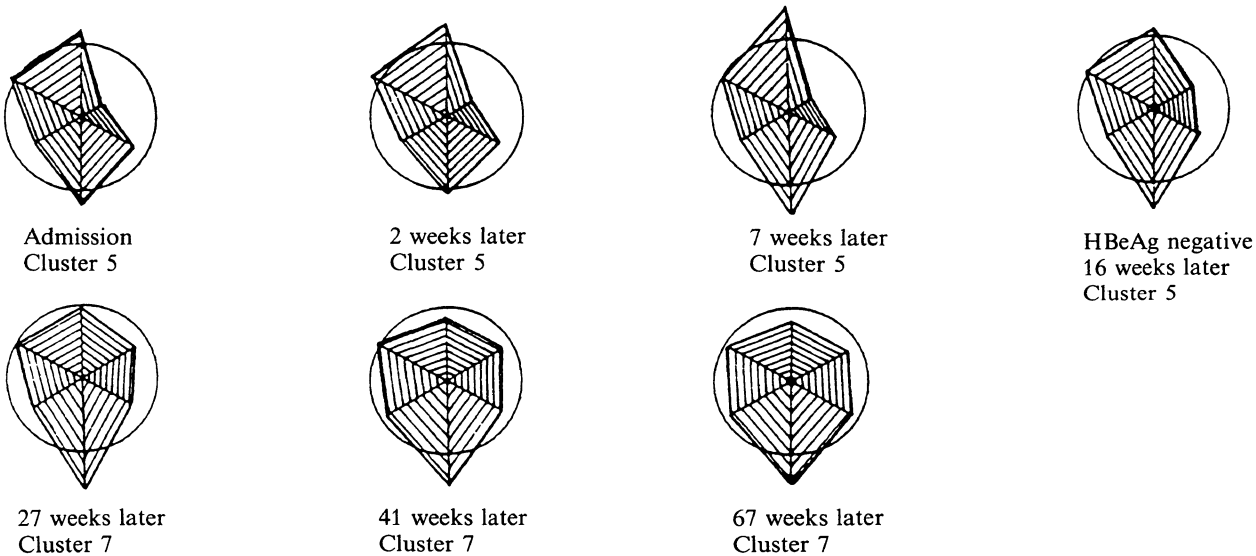
Cluster	Final Diagnoses																				n
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	25	3	3	5	2	2		7	4	10	13										74
2														1							1
3								3		2	1	2		3	1						12
4										2		1		1		1					5
5										12	1	1									14
6	4					2		3	2	10	6	1	3						1		32
7	5						1	12	17	3	6	1									45
8																	1				1
9	1								1	1								3	1		7
n	35	3	3	5	2	4	1	25	24	40	27	6	3	5	1	1	1	3	1	1	191

course of a disease is facilitated if the physician can watch the development of a pattern instead of looking for changes of single parameters.

Observations in 66 patients, who had been examined, diagnosed and allocated to a group of the model

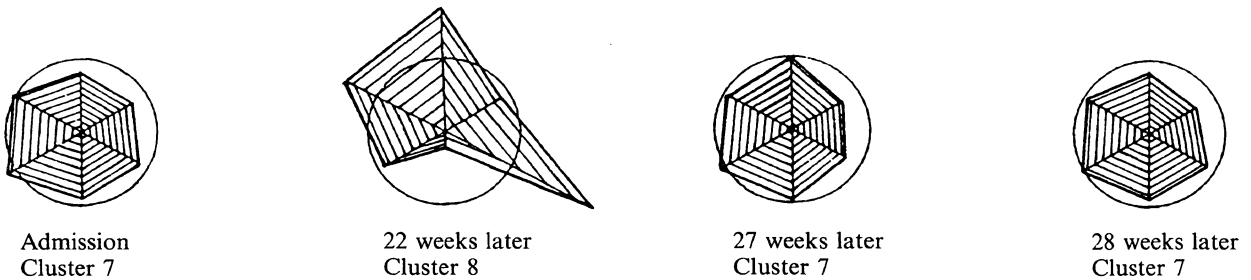
repeatedly at different times, showed that changes in diagnosis or in therapy are often correlated with change of the assigned cluster. By means of two representative cases, figure 6 demonstrates that the application of the model could be a useful tool in longitudinal supervision of patients with liver diseases.

Patient 1



Patient 1 suffered from a HBsAg and HBeAg positive chronic active hepatitis with cirrhosis, severe portal hypertension, splenomegaly and with history of ascites, gastrointestinal bleeding and encephalopathy and was assigned to cluster 5 at the point of admission. During a period of 67 weeks of supervision seroconversion from HBeAG pos to neg occurred, followed by continuous clinical improvement and remaining assignment to cluster 7.

Patient 2



Patient 2 had suffered from steatosis of liver for some years. At admission he was assigned to cluster 7. At the following examination an acute A hepatitis was diagnosed. From the results of the same date the patient was allocated to cluster 8, showing an impressive clinical chemical pattern. Some weeks later he returned to the group he had been attached to at the first examination.

Fig. 6. Cluster membership of patients during the course of disease.

It should be mentioned that the model is based on a certain data base and is only valid for this in a strict sense.

In conclusion, the application of cluster analysis based on clinical chemical parameters leads to a clinically meaningful classification of patients with liver diseases.

## References

1. Solberg, H. E., Skrede, S. & Blomhoff, J. P. (1975) Diagnosis of liver diseases by laboratory results and discriminant analysis. *Scand. J. Clin. Lab. Invest.* 35, 713–721.
2. Solberg, H. E., Skrede, S., Elgjo, K., Blomhoff, J. P. & Gjone, E. (1976) Classification of liver diseases by clinical chemical laboratory results and cluster analysis. *Scand. J. Clin. Lab. Invest.* 36, 81–85.
3. Winkel, P., Ramsoe, K., Lyngbye, J. & Tygstrup, N. (1975) Diagnostic value of routine liver tests. *Clin. Chem.* 21, 71–75.
4. Schmidt, E. & Schmidt, F. W. (1980) Anwendung von Bewertungsverfahren – Modell Leber-Erkrankungen – Klinik. In: *Validität klinisch-chemischer Befunde* (Lang, H., Rick, W., Büttner, H., eds.) Springer, Berlin, pp. 92–112.
5. Sandel, P. & Vogt, W. (1978) A comparison of discriminant methods. In: *Computing in Clinical Laboratories* (Siemaszko, F., ed.) Pitman Medical Publishing, Kent, pp. 272–282.
6. Vogt, W., Sandel, P., Schwarzfischer, P., Braun, S. L., Langfelder, Ch. & Knedel, M. (1981) Cluster-oriented discriminant analysis; taxonomic classification of thyroid function. *Clin. Chim. Acta* 112, 213–223.
7. Vogt, W., Sator, H. & Nagel, D. (1984) Computer assisted clinical chemical diagnosis, pattern cognition and recognition by cluster analysis. *Trends Anal. Chem.* 3, 166–171.
8. Vogt, W., Nagel, D. & Sator, H. (1987) Cluster analysis in clinical chemistry: A model. John Wiley and Sons, New York.
9. Späth, H. (1977) Cluster-Analyse-Algorithmen zur Objektklassifizierung und Datenreduktion, 2nd rev. edn. R. Oldenbourg, München.
10. Ward, J. H. (1963) Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58, 236–244.
11. Friedman, H. P. & Rubin, J. (1967) On some invariant criteria for grouping data. *J. Am. Stat. Assoc.* 62, 1159–1178.
12. Hubert, L. J. & Arabie, P. (1985) Comparing partitions. *J. Classif.* 2, 193–218.

Prof. Dr. W. Vogt  
Institut für Klinische Chemie  
und Laboratoriumsmedizin  
Deutsches Herzzentrum München  
des Freistaates Bayern  
Lothstraße 11  
D-8000 München 2